METHOD AND APPARATUS FOR IDENTIFYING A DIGITAL AUDIO SIGNAL

1. Field of the Invention

[0001] This application claims priority under 35 U.S.C. § 119 to U.S. Application No. 09/973,893, filed on October 11, 2001, the entire content of which is hereby incorporated herein by reference. The present invention relates to method and apparatus for identifying a program signal that is broadcast to members of an audience. More particularly, the present invention relates to method and apparatus for identifying a program signal having a digital audio component. Preferably, such method and apparatus will find use in audience measurement and/or broadcast monitoring services.

2. Related Art

[0002] Third party measurements are typically used in the broadcasting industry to verify that program elements (e.g., commercials) are disseminated in accordance with contractual arrangements, and to estimate the size and composition of the audience. One measurement technique involves reading an ancillary encoded identification label or signal that is transmitted with the program. Another measurement technique involves extracting characteristic features (commonly called "signatures") from the program, and then comparing the extracted features with a library of features from known program elements.

[0003] There is some degree of overlap in the code and signal processing between the above two techniques. If a broadcast program is encoded with a label associating it with a final distributor of the program (e.g., a local news broadcast labeled as such by the originating station), that label may be essentially self-attesting, or may be interpreted by

recourse to a master look-up table. On the other hand, if the encoded label only identifies an originator or intermediate distributor (e.g., a syndicated program labeled only with the program name and episode number), then it may be necessary to compare that label with a library of labels collected from a number of local stations, in order to identify the station that transmitted the program in question.

[0004] U.S. Patent No. 5,481,294 to Thomas et al. (incorporated herein by reference) discloses apparatus and method whereby a program signal that is to be identified is initially processed to extract an ancillary identifying code. If the code is found, it is stored along with the time at which it was received or otherwise selected by the user (known as the "read time") in a memory for subsequent transmission to a central data collection and processing facility. If no code is found, a signature is extracted from the program signal, stored as a time-stamped record in the memory, and subsequently communicated to the central facility where it is compared with similar signatures extracted from known programs at monitoring facilities that may be remote from the central facility.

[0005] A television measurement system of the sort taught by Thomas et al. may employ codes written into either the video or audio components of the composite broadcast signal, and also may extract signatures from either the video or audio portion of that signal. Moreover, one may configure a system of this sort to extract signatures from a program signal even when an identifying label or code is read. An arrangement of this sort can provide signature data to provide "fill-in" identification at times between sequential transmissions of an audio code.

[0006] Also of note is the identification tag reading system disclosed in U.S. Patent No. 6,202,218 to Ludtke (incorporated herein by reference). The disclosed measurement system is embedded in an in-home entertainment network having consumer electronic equipment adapted to communicate with other such equipment by means of an IEEE

1394 serial interface. Ludtke discloses an arrangement in which a program-identifying label received with a broadcast data stream that is used in the household is read, interpreted, and forwarded to a remote data collection entity.

[0007] The advent of digitally-transmitted television signals has had a profound impact on systems for both verifying broadcasts and for determining audience viewing preferences (in dwellings statistically selected to participate in a television audience measurement). For example, some video encoding arrangements that work well with analog video signals are incompatible with digital transmission because such codes, if present on an original analog signal, do not survive the data compression that is part of the process of converting the analog video signal to a digital one. Moreover, there are a variety of applicable digital transmission standards and a variation in the extent to which broadcasters adhere to those standards. For example, in 2001, the U.S. market faced several digital transmission standards: one for over-the-air terrestrial transmission, one for cable distribution, and one for satellite-to-end-user transmission.

[0008] Although the audio component of the overall program signal generally utilizes far less valuable bandwidth than does the video component, there is variation among digital audio standards. The ATSC (Advanced Television Systems Committee) standard, for example, mandates the use of what is called AC-3 audio, which could also be carried by direct satellite and cable systems. In 2001, the majority of satellites (Direct Broadcasting Satellite – DBS) and digital cable system were not using AC-3 sound. For example, some U.S. UBS used a standard referred to as Musicam or MPEG1 Layer 2 audio. Other standards, such as the Japanese AAC (Advanced Audio Coding) standard using MPEG2, Layers 1-3, are also known. Of course, many programs are still in a linear PCM format.

[0009] Generally speaking, the compressed audio formats call for each audio signal stream to be formatted into frames, where each frame can be configured as a string of packets that can be broadcast at a single frequency, or in a single channel with other

audio streams by means of time domain multiplexing. For example, MPEG breaks each frame into a set of fixed-duration packets (where each packet has a header) for multiplexing audio and video bitstreams in one stream with the necessary information to keep the streams synchronized when decoding. Each audio frame is autonomous and contains all the information necessary for decoding so that it can be processed independently of previous or subsequently transmitted frames. Although the length of a frame may vary, depending on the bit rate and sampling frequency, there is a maximum allowed frame length, and smaller frames (i.e., those arising from lower fidelity audio signals) may be padded with dummy data in order to provide a fixed interval between frame headers.

[0010] Many of the compressed audio standards set aside portions of selected packets for the transmission of auxiliary data (e.g., signal identification data) that is not part of the audio signal. Moreover, almost all such standards provide for the use of padding bytes in order to provide a fixed interval between frame headers. These padding bytes can, in some cases, be used for the purpose of adding an ancillary program identification label even if a defined auxiliary data field is not provided for in a given standard. Thus, it is expected that in at least the great majority of packetized digital audio broadcasting systems, a program-identifying label can be added to a predetermined portion of a packet or frame.

[0011] Consumer electronic equipment for receiving digital broadcasts typically have a standard digital audio output from the receiver to the consumer's digital audio equipment. The receiver may comprise a digital tuner in a set-top box that provides an analog output to an NTSC receiver, a digital television receiver and display, or any of a number of other known audio receivers. Moreover, the use of digital audio equipment has made it common for the consumer digital receiver to supply a digital audio output even if the input to the receiver is an analog broadcast signal or an input from an analog VCR. That is, the consumer's receiving apparatus (that is to be monitored) may be used

not only for receiving both analog and digital broadcast program signals, but also for selecting a program signal source from a number of possible local sources such as DVD or CD players.

[0012] The industry-standard design for providing a digital audio output signal for use by digital audio equipment is known as the Sony-Philips Digital Interface (SP/DIF). The signal available at an SP/DIF connector may be either a uncompressed linear PCM (Pulse Code Modulated) digital signal having a bit rate of no more than 64 kbit/sec, or may be a non-linear PCM encoded audio bitstreams, such as, in the AC-3 format and having a bit rate of 384 kbit/sec.

[0013] Thus, what is needed is a digital signal recognition system which is capable of accurately and reliably recognizing digital audio signals in all of the various configurations and implementations described above.

SUMMARY OF THE INVENTION

[0014] It is an object of the present invention to provide method and apparatus for processing received digital audio signals, transmitted through a wide variety of media, to ensure accurate recognition.

[0015] According to a first aspect of the present invention, digital audio signal recognition apparatus includes an input connector for direct connection to a standard SP/DIF output connector on consumer digital television receiving equipment. The apparatus includes structure for processing the input digital audio signal to obtain one or more of: (i) an identifying label encoded in a portion of a digital audio frame; (ii) an identifying label code embedded in a decompressed audio signal; (iii) a copy of a selected portion of a frame (e.g., a checksum portion); and (iv) a copy of a selected portion of the decompressed digital audio signal. Preferably, the apparatus also has an SP/DIF output connector and provides a repeated copy of the input digital audio signal

at that output, so that the user can use the audio and video equipment without interference by the signal recognition process.

[0016] According to another aspect of the present invention, a method of collecting tuning data from digitally transmitted program signals comprise an initial step of obtaining the digital audio signal associated with the program, using a SP/DIF connection. If that signal is a non-linear, PCM encoded audio bitstream signal, an attempt may be made to read a first type of program-identifying label from an auxiliary message portion of the digital signal or to select a predetermined selected portion of the signal frame as a first candidate signature. In addition, the input signal may be decompressed and the measuring equipment may attempt to read a second type of program-identifying label distributed as a code embedded in the audio signal or to generate a second candidate signature from a predetermined part of the decompressed audio signal. Of course, if the signal at the SP/DIF connection is an uncompressed, or a linear PCM, or already a decompressed digital audio signal, the second label and the second candidate signature can be collected without an intermediate decompression step. On the other hand, if the first label and the first candidate signature are present, they can be collected too.

[0017] According to yet another aspect of the present invention, a method of collecting tuning data includes the steps of obtaining an input digital audio signal at an SP/DIF connection, and processing the input signal in four parallel paths to obtain up to four identification data. The identification data may comprise (i) any program-identifying label that is present in the compressed audio signals, (ii) any program-identifying label that is present in the decompressed audio signals, (iii) a first candidate signature from the input signal (if it is in a compressed audio format), and (iv) a second candidate signature from the decompressed audio. All of the four possible identification data that are collected at each measurement time are assembled to form a time-stamped record that may be communicated to a store-and-forward apparatus. The store-and-forward apparatus stores in a memory at least some of the time-stamped records sent to it, and subsequently forwards the stored records to a central data collection facility. It will be

clear to those skilled in the audience-measurement arts that not all of the collected data (e.g., data collected during a time when the television is turned off, or data that supports a temporal resolution finer that what is called for in the measurement) is of value, and that the triage operation performed by the store-and-forward apparatus, or by some other portion of the measurement system installed a user dwelling, reduces the cost of storing and communicating data.

[0018] According to a further aspect of the present invention, a digital signal recognition system comprises monitoring equipment tuned to all the broadcast signal sources that can be viewed in statistically selected dwellings. Preferably, the monitoring equipment collects and stores all of the first and second program-identifying labels that are present in the signal, as well as extracting reference versions of both the first and second signatures. These data are stored as time-stamped records where the interval between data collection times is the smallest acceptable interval in any of the measurements for which the data are to be used. For example, if data from a monitoring site are to be used in both an audience measurement having a minimum reportable viewing interval of ten seconds and in a commercial broadcast verification study having a contractual resolution of one half second, the monitoring site will collect data with a one half second resolution. Regardless of the resolution involved, a central facility periodically retrieves data from both one or more monitoring sites and from some predetermined number of selected dwellings, and compares the various signatures and intermediate codes in order to identify the program signals selected by the sampled audience members.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] Figure 1 is a data structure diagram of a SP/DIF bitstream where AC-3 audio data, and ID label and signature for audience research are located or extracted.

[0020] Figure 2 is a system-level block diagram of a preferred measurement system according to an embodiment of the present invention.

[0021] Figure 3 is a block diagram of the measurement equipment depicted in Fig. 2.

[0022] Figure 4 is a logic diagram depicting a main loop of a program executed by a tuning measurement apparatus of the invention.

DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EXEMPLARY EMBODIMENTS

1. Introduction

[0023] While the present invention will be described with respect to television audience monitoring (e.g., Nielsen television rating) systems, it should be understood that the present invention applies equally well to radio audience monitoring, Internet audience monitoring, radio/TV commercial verification, copyright royalty collection, etc. As used herein, the term "program signal" refers to segments of various lengths such as all or parts of programs, commercials, promotional messages, public service announcements, and the like, as well as signals generated from consumer program signal storage equipment such as video cassette recorders (analog or digital), CD players, VCD players, DVD players and the like.

[0024] Briefly, the preferred embodiment receives a digital program signal and analyzes it in one or more of four different ways. First, the preferred embodiment can identify any program-identifying label that is present in compressed digital audio signals.

Second, the preferred embodiment can identify any program-identifying label that is present in decompressed audio signals. Third, the preferred embodiment can identify a first candidate signature from the input signal (preferably when it is in a frame format). And fourth, the preferred embodiment can identify a second candidate signature from the decompressed audio signals. One or more of these identifications are stored in a time-stamped record which may be used immediately, or at a later time, to verify transmitted information or measure audience participation.

2. The Signal Format

[0025] Fig. 1 depicts how AC-3 and Nielsen data fit within a SP/DIF bitstream. As depicted in Fig. 1, a SP/DIF bitstream signal 10 (as an example, here the audio is compressed in AC-3 format) is formatted based on IEC 60958 standard of International Electrotechnical Commission. It comprises a sequence of fixed-length frames 12. Each frame contains two sub-frames 13. Each of the sub-frame comprises a Preamble field 14, a Auxiliary data field 15, the data field 16 that is delivered by complying with IEC 61937 standard later, and the Status field 17. According to IEC 61937, a data burst is comprised of many IEC 60958 data fields 16. Data bursts are separated by Stuffing 20. The length of data burst is variable. Each data burst comprises a Header 21, Burst Information 22, Length 23, and a Payload 24 whose length is indicated by Length 23. Here the payload is carrying the AC-3 data. The Payload comprises SI 25, BSI 26, a number of Audio Blocks 27, and Aux data field 28, and finally, CRC2 29 which is the CRC for the entire AC-3 Frame. The last two sections are employed according to the present invention in order to conduct media research that includes measuring viewing information of the audience and monitoring programs.

[0026] The Stuffing 20 between data bursts is all zeros, however it must be noted that the IEC60958 frames still contain preamble, channel status bits, validity bits, etc. Bits 12-27 are zero. The stuffing between data bursts is used to maintain the proper synchronization of the audio out since the data channel has a capacity of a higher data

rate than is necessary to convey the compressed audio content. In the "worst case" scenario, where the capacity of the data rate is fully utilized, there would be no spare Stuffing space left. However, according to statistical data from real audio contents, there typically are Stuffings available that can be altered and used for delivering additional viewing activity information from the receiver to the outside, by a resident software meter preinstalled inside the receiver. The viewing activity information can be a detailed description of what the audience(viewer) is doing with the receiver and what the receiver is doing accordingly.

[0027] In the case of digital television, video and audio data in compressed form are carried in a bit-stream using a format specified by the Advanced Television Standards Committee (ATSC). The audio data uses Dolby's AC-3 compression algorithm and the AC-3 bit-stream contains, besides the actual audio data, headers containing additional information such as synchronization, timing and sampling rate. The AC-3 stream is organized into frames and each frame contains sufficient data to reconstruct audio corresponding approximately to the duration of video frame. Each frame has a fixed size in terms of total number of 16-bit words. At the end of a frame in addition to a Cyclic Redundancy Check (CRC2 in AC-3 terminology) word designed to detect errors in the reception of the frame, there is a reserved field for inserting auxiliary data (AUXDATA). Use may be made of the AUXDATA field to carry program and station information relevant to TV audience metering. Preceding AUXDATA are two fields: AUXDATAE is a 1-bit flag which indicates valid AUXDATA is present and AUXDATAL is a 14-bit field which indicates how many bits of auxiliary data are present.

[0028] In the case of 48 kHz-sampled audio with 16-bits per sample, each AC-3 frame represents, in a compressed form, 6 "blocks" of audio. Each block is derived from 256 samples per channel. The number of channels can vary between 1 in the case of monophonic audio to 6 for the case of "5.1 channel surround sound". A multi-pass

algorithm attempts to compress the data from each 256-sample block of audio in order to minimize the number of bits required to represent it. Since the frame size is fixed, at the end of the optimization process several bytes are usually available as "surplus" – in the bit stream these are defined by SKIPLE - if this bit is a "1" it means there are dummy bytes packed into the stream. Following this there is a 9-bit number SKIPL, which defines the number of bytes to skip at this point in the stream.

[0029] Most current AC-3 bit-stream generators do not make use of the auxiliary data field and as a result AUXDATAE is set to 0. In such cases, in order to utilize the AUXDATA feature, the stream may be modified by at first examining each frame to determine the total number of SKIP bytes present in the frame. These will occur at the end of each block. These will be used to create the necessary space for AUXDATA at the end of the frame. By appropriately modifying the SKIPL values at the end of each block and repacking the bits, a desired amount of space can be created for AUXDATA. In the event adequate bits to meet the AUXDATA requirements are not present, no change to the frame is made and AUXDATAE is set to 0. It may be noted that depending on the nature of the audio, not every frame will be capable of carrying auxiliary data. Indeed the Cyclic Redundancy Check words CRC1 and CRC2 have to be recomputed after these changes are made.

3. The Signal Identification Techniques

[0030] In view of the above discussion, a first technique for identifying a broadcast in a monitoring or audience-measurement system having a clock or other time keeping means operatively associated therewith includes the steps of reading a program-identifying label from an auxiliary data field of a digital audio signal frame 40, and associating the label with the time at which it was read as a time-stamped record that can be stored in a memory for subsequent communication to a central data processing facility. The second technique is similar to the first, but obtains the program-identifying label after decompressing a compressed digital audio signal.

[0031] In the first and second techniques, the number of bytes required for a program-identifying label can be relatively small. Thus, it may be possible to add a code to a signal frame even if no auxiliary data field is provided, or if the provided field is pre-empted for some other use. One could encode a program signal by writing the code in non-used portions of the frame – e.g., in padding bytes that are otherwise ignored by the ordinary audio data processing operations.

[0032] A third technique for a tuned program signal is to extract a characteristic feature, or signature, from the signal at both a statistically selected tuning site and at one or more monitoring site(s) arranged so as to monitor all broadcast signals that can be received at the tuning site. The candidate signature from the tuning site can then be compared to reference signatures from the monitoring site or sites in order to identify the tuned program by matching the signatures. Correspondingly, the broadcast of repeated program elements can be identified by comparing candidate signatures from a monitoring site with a library of reference signatures. The fourth technique is similar to the third technique, but extracts the candidate signature after decompressing a compressed digital audio signal, or from original uncompressed digital signal.

[0033] In the third and fourth techniques, the well-defined data formats used for the transmission of digital signals facilitates the identification of broadcast programs by comparing signatures. One way to compare signatures is to extract a predetermined field from a frame 10 of a digital audio signal at both a measurement site (tuning site and/or monitoring site) and a reference site. In a preferred embodiment of the invention a checksum field (which is commonly a CRC checksum) is read from each digital audio frame and is associated with a read time output from a clock or other time keeping means in a step that forms all or part of a time-stamped record. The CRC is a desirable signature because it comprises a relatively small data field that is variable enough to yield a unique signature. It will be recognized, however, that many other portions of a

data frame 10 (e.g., the non-program labeling contents of an auxiliary data field) could equally well be used for this purpose.

[0034] The comparison of signatures extracted from analog signals usually requires circuitry or signal processing for handling both (i) temporal errors or drifts, and (ii) changes in the magnitudes of the signals acquired at two different locations. In the case of digital signals, however, the recognition or matching process is considerably simpler. Although the identifying algorithms must provide for "sliding" data blocks relative to each other along a time axis in order to accommodate temporal drifts or other time-keeping errors, there is no corresponding signal amplitude problems inasmuch as two matching signals will be substantially identical bit for bit.

[0035] The preferred embodiment of the invention combines all four techniques in reading a program-identifying label from an auxiliary data field of a digital audio signal frame, in reading a program-identifying label of embedded code from a decompressed or uncompressed audio signal, in reading a predetermined portion of the signal frame, in extracting a signature from a decompressed or non compressed audio signal. The labels, if found, and the predetermined portion of the signal frame and the signature are associated with the local time when the frame was read in order to generate a time-stamped tuning record for each frame of digital audio signal that is received. During operation, there may not be all four techniques present at the same time. For example, the original audio signal may be an uncompressed (linear) digital audio signal. Then, only the third and fourth techniques may be employed. Even in a compressed digital signal case, there may be a lack of a program label in an auxiliary data field for various reasons. As long as the system gets at least one label or one signature, they (it) will be associated with the local time when the frame was read in order to generate a time-stamped tuning record for each frame of digital audio signal that is received.

4. The Structure

[0036] In a preferred embodiment depicted in Fig. 2, a tuning or verification site 36 comprises a receiver 34, a tuning measurement apparatus 48, a clock or timing device 38, and a storage and forwarding apparatus 52. The tuning measurement apparatus 48 receives a digital audio signal output by a consumer's receiving hardware 34 at an industry-standard SP/DIF connector 50. The tuning measurement apparatus 48 decodes the signal to read a program-identifying label, if any, and to collect a predetermined portion of each signal frame for signature analysis (to be described below). The tuning measurement apparatus 48, the clock or timing device 38, and the storage and forwarding apparatus 52 (Fig. 2) may be embodied in a single computer, or in a plurality of processors, or in hard-wired circuitry. These circuits may also be incorporated into the dwelling set-top box, or as a stand-alone device.

[0037] The time-stamped records that are generated by the tuning measurement apparatus 48 are sent to a data storage and forwarding apparatus 52 that stores all or some subset of these records in a memory 44, for subsequent transmission over a public switched telephone network 54 to a central data collection facility 46 by means of a modem 53. It will be recognized that instead of a dial-up modem, other suitable communication means such as a cable modem, or a wireless data link could be used for this purpose.

[0038] In more detail, the tuning measurement apparatus 48 is configured to be connected to a SP/DIF connector 50 that is part of a consumer-owned receiving apparatus 34, which may be a digital television receiver, a set-top box feeding an analog signal to an NTSC receiver, or any other such receiving apparatus. As is provided for in industry standards, the signal at the SP/DIF connector may be either a linear PCM (uncompressed) digital audio signal at 64 kbit/sec or less, or a non-linear PCM encoded audio bitstream signal in the AC-3 format and having a bit rate of 384 kbit/sec. Although the standard allows for multiplexed AC-3 streams, the equipment that was initially placed on the market transmitted only one stream. If two or more streams were

transmitted, the apparatus of the invention may be configured to select one of the AC-3 streams for decoding. It may be noted for any signal appearing at the SP/DIF connector, the apparatus 48 does not need to deal with the full DTV bitstream, which has a much higher bit rate of 19.2 Mbit/sec. Hence, the preferred embodiment is expected to be both less expensive and more reliable than alternate apparatus that acquires the full DTV signal from within the consumer's receiving equipment.

[0039] The records not stored in the memory 44 may be discarded in the interest of using the memory 44 efficiently while still providing the measurement's specified temporal signal granularity. For example, if an audience measurement is made with a guaranteed resolution of fifteen seconds from a digital broadcast signal having an interval of eight milliseconds between frame headers, the memory 44 need only retain one out of every 1875 records taken while the receiving equipment is in active use. Of course, no data need be collected when the equipment is not in use. Hence if the monitored SP/DIF connector always provides an output (e.g., as might be the case if an always-on set-top box digital receiver is used to provide a signal to an NTSC television), a separate on/off sensor 56 can provide an input to the storage and forwarding unit to indicate the time periods during which data are to be collected.

[0040] The store-and-forward apparatus 52 in the preferred embodiment assembles time-stamped records from the inputs. As noted previously, the amount of data sent from the measurement apparatus 48, which is preferably taken from each frame of the audio signal, is far in excess of what is required for an audience measurement. Hence, the store-and-forward apparatus 52 may filter the inputs and generate time-stamped records from only some of them. A monitoring site used for measurement, on the other hand, may use essentially the same processes to collect and save all the available data for subsequent comparison with data from a plurality of statistically selected dwellings.

[0041] As shown in Fig. 3, in order to avoid attenuating the signal at the receiver output, the apparatus 48 preferably comprises a repeater circuit 64 that provides a repeated signal to a second SP/DIF connector 66 that can be used to furnish signals to other consumer-owned equipment. The signal input to the apparatus 48 is fed to a bit rate detector 68 used to determine whether the signal is a uncompressed (linear) PCM signal or an AC-3 signal, to route the AC-3 signals to a frame decoder 70, and to route uncompressed PCM signals to both a embedded code reader circuit 72 (designed to extract embedded audio codes from the signal) and to a signature extractor circuit 74. If the signal is in the AC-3 format, the frame decoder 70 provides the contents of an auxiliary data field (if one is present) and a copy of a predetermined portion of the frame (e.g., the CRC field) as outputs to a store-and-forward apparatus 52. The frame decoder 70 also has an output to a decompression circuit 76 that supplies a digital signal to the embedded code reader 72 and the signature extractor 74. The signature extractor 74 may supply the candidate signature from the decompressed signal, or the candidate signal from the uncompressed signal to the storage and forwarding apparatus 52.

[0042] The preferred embodiment may collect data from a linear PCM audio signal, which may be originally uncompressed, or obtained from the transmitted non-linear PCM encoded audio bitstream signal by a decompression process that is part of the standard operation of recovering the signal that was compressed prior to transmission. The third and fourth techniques mentioned previously can be employed in these kinds of situations.

[0043] U.S. Patent Application Serial No. 09/116,397, filed July 16, 1998 and assigned to the assignee of this application, and U.S. Patent Application Serial No. 09/428,425, filed October 27, 1999, and U.S. Patent Application Serial No. 09/543,480, filed April 6, 2000 (each of which is incorporated herein by reference) disclose methods and apparatus for encoding audio signals by spectral modulation. These coding arrangements are selected so that the code survives subsequent compression and

decompression and is hence compatible with various digital signal transmission standards. It will be recognized that other coding arrangements that have been (and will be) developed satisfy this process. Hence, the preferred embodiment of the invention attempts to recover an encoded program label from the PCM audio signal, which may be a decompressed audio signal. In other arrangements, of course, audio codes may be recovered from an analog audio signal, such as one recovered from a microphone adjacent a speaker.

[0044] As discussed earlier, it is known in the broadcast measurement arts to extract signatures from video and/or audio signals and to compare these with reference signatures extracted by similar means from known program signals. In the preferred embodiment of the invention, a signature is extracted from a linear PCM audio signal, and is stored, with other identification data as a time-stamped tuning record. Although the preferred arrangement calls for extracting this signature from a PCM digital audio data stream, those skilled in the art will recognize that one could also elect to convert the digital audio signal to an analog signal and then extract the signature by methods such as those described in U.S. Patent Nos. 4,697,209 and 4,677,466, (both of which are incorporated herein by reference).

5. The Process

[0045] The preferred embodiment of the invention runs four identification processes on a received signal and generally operates in a parallel fashion, as depicted in the flow chart of Fig. 4. It will be recognized that inasmuch as any one of those processes can yield a positive identification of a program signal, fewer identification processes can also be used. Moreover, instead of collecting all of the identification data all of the time, one could elect to set up a hierarchical collection scheme, such as the one disclosed by Thomas et al. in U.S. Patent No. 5,481,294 (incorporated herein by reference), that initially looks for a preferred identification datum and collects an

alternate datum only if the first one is not available. For example, one could configure a system in which a program label was initially sought in an AUX field 31 of a digital signal frame and in which an embedded code was sought in the decompressed digital audio signal only if the attempt to read the AUX field failed.

[0046] Moreover, one could also elect to run multiple identification processes and to then select from among the available data at the time a time-stamped record was assembled (step 42). For example, if a positive identification label was read from an AUX field, that code and the CRC signature could be saved as the time-stamped record. To continue with the example, if no AUX code was found but an embedded code was read, the CRC signature and the embedded code could be saved for retrieval by the central data collection facility.

[0047] In Fig. 4, the digital audio signal associated with the program is obtained at the SP/DIF connection in step S1. The program-identifying label is read from the auxiliary data field of a digital audio signal frame in step S2. A checksum field (which is commonly a CRC checksum) is read from each digital audio frame and is associated with a read time output from a clock or other time keeping means 38 in a step S3 that forms all or part of a time-stamped record. If the received digital audio signal is a compressed signal, it is decompressed at step S4. In step S5, an encoded program label is recovered from the decompressed audio signal, PCM audio signal. In step S6, a signature is extracted from a linear PCM audio signal, and is stored, with any other identification data, as a time-stamped tuning record in step S7. While the preferred embodiment conducts four signal identification procedures in parallel, any combination of two or more of these procedures could be conducted in parallel or in series. For example, the process could first determine whether the received signal was compressed, and then invoke the appropriate program-identifying and signature extraction steps. One could also elect to run any combination of these identification procedures, and then

select from among the available data at the time the time-stamped record was assembled.

[0048] As mentioned previously, the program label (e.g., identifying label) in the aux data field may be encoded by the broadcasters and/or program producers. As an alternative program identifying code, the broadcasters and/or any other participant in the distribution system may also insert PSIP (Program System Information Protocol), and/or Content Identification (Content ID) data, and/or
other useful data in the aux data field. As yet another alternative program identifying code, the broadcasters and/or other participant in the distribution system may also copy program related data from bitstream data areas outside the aux data field into the aux data field. The statistical availability of Stuffing can be also utilized to deliver viewing activity information. Furthermore, the SP/DIF standard allows for non-audio data instead of, or in addition to, the non-linear PCM embedded audio bitstreams. The digital signatures and auxiliary codes may be effectively applied to identify or monitor non-audio data that includes Internet and other data transmission applications. All such alternatives are within the scope of the appended claims.

6. Conclusion

[0049] Thus, what has been described is a digital audio signal recognition system and method which accurately and reliably detects digital audio signals. The system according to the present invention will find use in verification sites, reference library sites, audience monitoring dwellings, and in any site where the monitoring, storing, and/or comparison of digital audio signals is required.

[0050] The individual components shown in block or schematic form in the Drawings are all well-known in the signal processing arts or are described in the documents incorporated herein by reference, and their specific construction and/or operation are not critical to the operation or best mode for carrying out the invention.

[0051] While the present invention has been described with respect to what is presently considered to be the preferred embodiments, it is to be understood that the invention is not limited to the disclosed embodiments. To the contrary, the invention is intended to cover various modifications and equivalent structures and functions included within the spirit and scope of the appended claims.

WHAT IS CLAIMED IS:

1. Apparatus for identifying a digital audio signal, comprising:

an input which receives the digital audio signal; and

structure configured to (i) identify a program-identifying code in the received digital audio signal, (ii) identify a program-identifying code in a decompressed or uncompressed received digital audio signal, (iii) identify a feature signature in the received digital audio signal, and (iv) identify a feature signature in the decompressed or uncompressed received digital audio signal.

2. Apparatus according to Claim 1, wherein said structure time-stamps the identified signals.

3. Apparatus according to Claim 2, further comprising a clock which supplies a clock signal to said structure.

4. Apparatus according to Claim 2, further comprising a memory for storing the time-stamped identified signals.

5. Apparatus according to Claim 4, further comprising transmission structure for transmitting the stored, time-stamped identified signals to a collection site.

6. Apparatus according to Claim 1, wherein said structure is disposed in a dwelling.

7. Apparatus according to Claim 1, wherein said structure is disposed in a verification site.

8. Apparatus according to Claim 1, wherein said structure is disposed in a monitoring site.

9. Apparatus according to Claim 1, further comprising a tuner which receives a broadcast program signal and outputs the digital audio signal to said input.

10. Apparatus according to Claim 1, wherein the program-identifying code is placed in an aux data field in the digital audio signal by a broadcaster and/or a program producer.

11. Apparatus according to Claim 10, wherein the program identifying code is Program System Information Protocol (PSIP) data, Content ID data, and/or other useful data which are placed in the aux data field by the broadcaster and/or any other participant in the distribution system.

12. Apparatus according to Claim 10, wherein the program identifying code is program related data which are copied from bitstream data areas outside the aux data field into the aux data field by the broadcaster and/or other participant in the distribution system.

13. Apparatus according to Claim 10, wherein the Stuffing is altered with data which corresponds to viewing activity information.

14. Apparatus according to Claim 10, wherein the program-identifying code and the feature signal correspond to an Internet transmission.

15. Signal identification structure which receives an input digital audio signal, comprising:

signal processing structure configured to perform at least one of: (i) determining an identification code from a portion of a compressed received input digital audio signal, and extracting a signature from a portion of said compressed received input digital audio signal; and (ii) determining an identification code from a portion of an uncompressed received input digital audio signal, and extracting a signature from a portion of said uncompressed received input digital audio signal.

16. Structure according to Claim 15, wherein said signal processing structure performs both (i) and (ii).

17. Apparatus according to Claim 15, wherein said processing structure time-stamps determined and/or extracted signals.

18. Apparatus according to Claim 17, further comprising a clock which supplies a clock signal to said processing structure.

19. Apparatus according to Claim 17, further comprising a memory for storing the time-stamped signals.

20. Apparatus according to Claim 19, further comprising transmission structure for transmitting the stored, time-stamped signals to a collection site.

21. Apparatus according to Claim 15, wherein said processing structure is disposed in a dwelling.

22. Apparatus according to Claim 15, wherein said processing structure is disposed in a verification site.

23. Apparatus according to Claim 15, wherein said processing structure is disposed in a monitoring site.

24. Apparatus according to Claim 15, further comprising a tuner which receives a broadcast program signal and outputs the digital audio signal to said processing structure.

25. A method for identifying a digital audio signal, comprising the steps of:

receiving the digital audio signal at an input; and

configuring processing structure to (i) identify a program-identifying code in the received digital audio signal, (ii) identify a program-identifying code in a decompressed received digital audio signal, (iii) identify a feature signature in the received digital audio signal, and (iv) identify a feature signature in the decompressed received digital audio signal.

26. A method according to Claim 25, wherein said processing structure time-stamps the identified signals.

27. A method according to Claim 26, further comprising the step of supplying a clock signal to said processing structure.

28. A method according to Claim 26, further comprising the step of storing the time-stamped identified signals.

29. A method according to Claim 28, further comprising the step of transmitting the stored, time-stamped identified signals to a collection site.

30. A method according to Claim 25, wherein said processing structure is disposed in a dwelling.

31. A method according to Claim 25, wherein said processing structure is disposed in a verification site.

32. A method according to Claim 25, wherein said processing structure is disposed in a monitoring site.

33. A method according to Claim 25, further comprising the step of using a tuner to receive a broadcast program signal and output the digital audio signal to said input.

34. A signal identification process which receives an input digital audio signal, comprising the steps of:

performing at least one of:

i) determining an identification code from a portion of a compressed received input digital audio signal, and extracting a signature from a portion of said compressed received input digital audio signal; and

ii) determining an identification code from a portion of an uncompressed received input digital audio signal, and extracting a signature from a portion of said uncompressed received input digital audio signal.

35. A process according to Claim 34, wherein both (i) and (ii) are performed.

36. A process according to Claim 34, further comprising the step of time-stamping determined and/or extracted signals.

37. A process according to Claim 36, further comprising the step of supplying a clock signal to at least one of said determining steps.

38. A process according to Claim 36, further comprising the step of storing the time-stamped signals.

39. A process according to Claim 38, further comprising the step of transmitting the stored, time-stamped signals to a collection site.

40. A process according to Claim 35, wherein said determining step is performed in a dwelling.

41. A process according to Claim 35, wherein said determining step is performed in a verification site.

42. A process according to Claim 35, wherein said determining step is performed in a monitoring site.

43. A process according to Claim 35, further comprising the step of using a tuner to receive a broadcast program signal and output the digital audio signal to said determining step.

44. A process according to Claim 35, further comprising the step of decompressing the received digital audio signal.

45. A computer readable storage medium which causes one or more computers to carry out a method for identifying a digital audio signal, the computer readable storage medium causing the one or more computers to perform the steps of:

inputting the digital audio signal; and

configuring processing structure to (i) identify a program-identifying code in the received digital audio signal, (ii) identify a program-identifying code in a decompressed received digital audio signal, (iii) identify a feature signature in the

received digital audio signal, and (iv) identify a feature signature in the decompressed received digital audio signal.

46. A computer readable storage medium according to Claim 45, wherein said processing structure is configured to time-stamp the identified signals.

47. A computer readable storage medium according to Claim 46, further comprising the step of supplying a clock signal to said processing structure.

48. A computer readable storage medium according to Claim 46, further comprising the step of storing the time-stamped identified signals.

49. A computer readable storage medium according to Claim 48, further comprising the step of transmitting the stored, time-stamped identified signals to a collection site.

50. A computer readable storage medium according to Claim 45, wherein said processing structure is disposed in a dwelling.

51. A computer readable storage medium according to Claim 45, wherein said processing structure is disposed in a verification site.

52. A computer readable storage medium according to Claim 45, wherein said processing structure is disposed in a monitoring site.

53. A computer readable storage medium according to Claim 45, further comprising the step of using a tuner to receive a broadcast program signal and output the digital audio signal to said processing structure.

54. A computer readable storage medium which causes one or more computers to carry out a signal identification process which receives an input digital audio signal, the computer readable storage medium causing the one or more computers to perform the steps:

performing at least one of:

i) determining an identification code from a portion of a compressed received input digital audio signal, and extracting a signature from a portion of said compressed received input digital audio signal; and

ii) determining an identification code from a portion of an uncompressed received input digital audio signal, and extracting a signature from a portion of said uncompressed received input digital audio signal.

55. A computer readable storage medium according to Claim 54, wherein both (i) and (ii) are performed.

56. A measurement apparatus for determining which of a plurality of digitally broadcast program signals is tuned by a broadcast receiver apparatus having a receiver output connector for outputting a digital audio signal, the measurement apparatus comprising:

an input connector configured to connect the measurement apparatus to the receiver output connector;

a repeater circuit configured to receive the digital audio signal and to supply a repeated copy thereof to an output connector portion of the measurement apparatus;

a clock configured to supply a read time output;

a decoding circuit having an input from the clock, the decoding circuit configured to read an ancillary code from the digital audio signal and to associate a read time therewith;

a signature extractor having an input from the clock, the signature extractor configured to extract a digital signature from the digital audio signal and to associate the read time therewith;

a memory configured to store a plurality of data records, each of the data records comprising a respective read time and at least one of the ancillary code associated therewith and the digital signature associated therewith; and

a communication device configured to communicate the plurality of data records to a data collection facility.

57. The measurement apparatus of Claim 56, wherein the digital signal comprises a sequence of frames, each of the frames comprising a respective auxiliary data field configured to contain the ancillary code, each of the frames further comprising a predetermined portion usable as the digital signature.

58. The measurement apparatus of Claim 56, further comprising:

a signal decompression circuit for converting the sequence of frames into a PCM digital signal;

a second decoding circuit configured to receive the PCM signal, to extract an embedded second ancillary code therefrom, and to associate the read time therewith;

a second signature extractor configured to receive the PCM signal, to extract a second digital signature therefrom and to associate the read time therewith; and

wherein each of the data records further comprises at least one of the embedded code and the second digital signature.

59. The measurement apparatus of Claim 56, wherein the digital signal comprises a PCM signal having the ancillary code embedded therein.

60. The measurement apparatus of Claim 56, further comprising structure which determines whether the digital audio signal output from the receiver is anon-linear PCM encoded audio bitstream signal or the linear PCM audio signal.

61. The measurement apparatus of Claim 56, wherein the receiver output connector comprises a SP/DIF connector.

62. A method for generating a time-stamped tuning record for determining which one of a plurality of program signals is selected at a receiving site by a receiver apparatus having a receiver output connector configured to output a digital audio signal uniquely associated with the selected one of the program signals, the digital audio signal comprising a sequence of frames, the method comprising the steps of:

receiving the digital audio signal from the receiver output connector;

supplying a repeated copy of the digital audio signal to an output connector other than the receiver output connector;

attempting to read, from a predetermined field of one of the frames, a respective first ancillary code from the digital audio signal;

copying a predetermined portion of one of the frames as a respective first digital signature;

assembling the time-stamped tuning record from a read time datum representing a time at which the one of the frames was received from the receiver output connector and at least one of the first ancillary code and the first digital signature; and

communicating the time-stamped record to a data collection facility.

63. The method of Claim 62, comprising the additional steps of:

decompressing anon-linear PCM encoded audio bitstream signal to yield a linear PCM signal;

attempting to read, from the linear PCM signal, a second ancillary code embedded therein; and

extracting a second digital signature from the linear PCM signal;

wherein the step of assembling the time-stamped record comprises including at least one of the second ancillary code and the second digital signature therein.

64. The method of Claim 62, wherein the predetermined portion copied as the first digital signature comprises a checksum field.

65. The method of Claim 62, wherein the predetermined field comprises an auxiliary data field.

66. The method of Claim 62, wherein the digital audio signal comprises a signal having an AC-3 format.

67. A method for generating a time-stamped tuning record for determining which one of a plurality of program signals is selected at a receiving site at a read time by a receiver apparatus having a receiver output connector configured to output a digital audio signal, the digital audio signal uniquely associated with the selected one of the program signals, the method comprising the steps of:

receiving the digital audio signal from the receiver output connector;

attempting to read an ancillary code embedded in the digital audio signal;

extracting a digital signature from the digital audio signal;

assembling the time-stamped tuning record from the read time and at least one of the ancillary code and first digital signature; and

communicating the time-stamped record to a data collection facility.

68. The method of Claim 67, further comprising a step of outputting a repeated copy of the digital audio signal.

69. The method of Claim 67, wherein the selected program signal comprises a broadcast program signal.

70. The method of Claim 67, wherein the selected program signal is generated from consumer program signal storage equipment disposed with the receiver apparatus in a dwelling.

71. A measurement apparatus for determining which of a plurality of program signals is selected by a receiver apparatus having an SP/DIF output connector for outputting a receiver output digital audio signal uniquely associated with the selected program signal, the receiver output digital audio signal having one of a non-linear PCM encoded audio bitstream signal format and a linear PCM audio signal format, the measurement apparatus comprising:

an input connector configured to connect the measurement apparatus to the SP/DIF output connector;

a clock configured to supply a read time output;

a first decoding circuit configured to receive the read time output, the first decoding circuit configured to read an ancillary code from the digital audio signal if the digital audio signal has the non-linear PCM encoded audio bitstream format and to associate a read time therewith;

a first signature extractor configured to receive the read time output, the first signature extractor configured to extract a first digital signature from the digital audio signal if the digital audio signal has the non-linear PCM encoded audio bitstream format and to associate the read time therewith;

a signal decompression circuit configured to convert the receiver output digital audio signal to a converted linear PCM audio signal if the digital audio signal has the non-linear PCM encoded audio bitstream format;

a second decoding circuit configured to receive the read time output, the second decoding circuit configured to read an embedded code from one of the receiver

output PCM digital audio signal and the converted linear PCM audio signal and to associate the read time therewith;

a second signature extractor configured to receive the read time output, the second signature extractor configured to extract a second digital signature from one of the receiver output PCM digital audio signal and the converted PCM digital audio signal, and to associate the read time therewith;

a memory configured to store a plurality of data records, each of the data records comprising a respective read time and at least one of the respective first ancillary code, the respective first digital signature, the respective embedded code and the respective second digital signature associated therewith; and

communication means configured to communicate the plurality of data records to a data collection facility.

72. The measurement apparatus of Claim 71, wherein the first digital signature comprises a copy of a predetermined portion of the non-linear PCM encoded audio bitstream signal.

73. The measurement apparatus of Claim 72, wherein the first decoding circuit is configured to read the ancillary code from an auxiliary data field portion of the non-linear PCM encoded audio bitstream signal.

74. The measurement apparatus of Claim 72, further comprising a repeater circuit configured to receive the receiver output digital audio signal and to supply a repeated copy thereof to an output connector portion of the measurement apparatus.

75. The measurement apparatus of Claim 72, wherein the communication means comprises a modem portion of a data storage and forwarding

apparatus and wherein the data storage and forwarding apparatus comprises the memory configured to store the plurality of data records.

76. A method for generating a time-stamped tuning record for determining which of a plurality of program signals is selected at a receiving site by a receiver having a receiver output connector configured to output a receiver output digital audio signal in either a linear PCM format or a non-linear PCM encoded audio bitstream format, the receiver output digital audio signal uniquely associated with a selected one of the program signals, the method comprising the steps of:

receiving the receiver output digital audio signal from the receiver output connector;

determining if the receiver output digital audio signal is the non-linear PCM encoded audio bitstream signal or the linear PCM signal;

if the receiver output digital signal is the non-linear PCM encoded audio bitstream signal, attempting to read, from a predetermined field of each frame, a first ancillary code from the receiver output digital signal;

if the receiver output digital signal is a non-linear PCM encoded audio bitstream signal, copying a predetermined portion of each frame as a first digital signature;

if the receiver output digital signal is a non-linear PCM encoded audio bitstream signal, decompressing the non-linear PCM encoded audio bitstream signal to yield a decompressed PCM digital signal;

attempting to read, from one of the receiver output PCM digital signal and the decompressed PCM digital signal, a second ancillary code embedded therein;

extracting a second digital signature from one of the receiver output PCM digital signal and the decompressed PCM digital signal;

assembling the time-stamped tuning record from a read time datum and at least one of the first ancillary code, the first digital signature, the second ancillary code and the second digital signature; and

communicating the time-stamped record to a data collection facility.

77. The method of Claim 76, further comprising a step of comparing, at the data collection facility, data from the time-stamped record with corresponding data from at least one of a plurality of time-stamped reference tuning records, each of the reference tuning records comprising a copy of the predetermined portion of each frame of a digital broadcast signal broadcast from a source receivable at the receiving site and a reference digital signature extracted from a decompressed PCM digital signal obtained from the digital broadcast signal.

78. The method of Claim 76, further comprising a step of repeating the received output digital audio signal at an output connector other than the receiver output connector.

79. The method of Claim 76, further comprising a step after the step of assembling the tuning record and before the step of communicating the tuning record to a data collection facility of storing the tuning record in a memory portion of a data storage and forwarding apparatus.

80. A broadcast audience measurement method for determining which one of a predetermined number of broadcast signals is tuned by a receiving apparatus at a statistically selected tuning site at a plurality of tuning times, wherein each of the broadcast signals comprises a respective digital audio component formatted as a respective sequence of data frames, and wherein the receiving apparatus comprises an output connector supplying the digital audio signal, the method comprising the steps of:

selecting a subset of the data frames from the tuned broadcast signal;

associating a respective tuning time with each of the data frames in the subset;

storing a set of tuning records in a tuning memory at the tuning site, each of the tuning records comprising a copy of a predetermined portion of a respective data frame in the subset and the respective tuning time;

communicating the set of tuning records to a central data collection facility;

receiving the predetermined number of broadcast signals at at least one reference site, and storing a respective set of time-stamped reference records in a reference memory, each of the sets of reference records comprising a copy of the predetermined portion of each of the data frames of the digital audio component associated with a respective one of the broadcast signals;

communicating the sets of reference records to the central data collection facility; and

comparing, at the central data collection facility, the set of tuning records with ones of the sets of reference records to determine which of the broadcast signals is the tuned broadcast signal.

81. The method of Claim 80, wherein at least one of the tuning records comprises an ancillary code read from an auxiliary data field portion of the respective at least one of the frames.

82. The method according to Claim 80, wherein the auxiliary data field portion includes skip bits rather than auxiliary data.